

This article was downloaded by: [EBSCOHost EJS Content Distribution - Superseded by 916427733]

On: 2 October 2010

Access details: Access Details: [subscription number 911724993]

Publisher Psychology Press

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## European Journal of Work and Organizational Psychology

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713684945>

### "I don't see me like you see me, but is that a problem?" Cultural influences on rating discrepancy in 360-degree feedback instruments

Regina Eckert<sup>a</sup>; Bjørn Z. Ekelund<sup>bc</sup>; William A. Gentry<sup>d</sup>; Jeremy F. Dawson<sup>e</sup>

<sup>a</sup> Center for Creative Leadership, Brussels, Belgium <sup>b</sup> Human Factors AS, Oslo <sup>c</sup> Agder University, Agder, Norway <sup>d</sup> Center for Creative Leadership, Greensboro, NC, USA <sup>e</sup> Department of Work & Organisational Psychology, Aston University, Birmingham, UK

First published on: 12 March 2009

**To cite this Article** Eckert, Regina, Ekelund, Bjørn Z., Gentry, William A. and Dawson, Jeremy F. (2010) "I don't see me like you see me, but is that a problem?" Cultural influences on rating discrepancy in 360-degree feedback instruments', *European Journal of Work and Organizational Psychology*, 19: 3, 259 – 278, First published on: 12 March 2009 (iFirst)

**To link to this Article:** DOI: 10.1080/13594320802678414

**URL:** <http://dx.doi.org/10.1080/13594320802678414>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

# **“I don’t see me like you see me, but is that a problem?”**

## **Cultural influences on rating discrepancy in 360-degree feedback instruments**

Regina Eckert

*Center for Creative Leadership, Brussels, Belgium*

Bjørn Z. Ekelund

*Human Factors AS, Oslo, and Agder University, Agder, Norway*

William A. Gentry

*Center for Creative Leadership, Greensboro, NC, USA*

Jeremy F. Dawson

*Department of Work & Organisational Psychology, Aston University,  
Birmingham, UK*

360-degree feedback from a variety of rater sources yields important information about leaders’ styles, strengths and weaknesses for development. Results where observer ratings are discrepant (i.e., different) from self-ratings are often seen as indicators of problematic leadership relationships, skills, or lack of self-awareness. Yet research into the antecedents of such self–observer rating discrepancy suggests the presence of systematic influences, such as cultural values. The present study investigates the variation of rating discrepancies on three leadership skills (decision making, leading employees, and composure) in dependence of one exemplary culture dimension (power distance) on data from 31 countries using multilevel structural equation modelling. Results show that cultural values indeed predict self–observer rating discrepancies. Thus, systemic and contextual influences such as culture need to be taken into consideration when

---

Correspondence should be addressed to Regina Eckert, Research Associate, Center for Creative Leadership, Avenue de Tervueren 270, 1150 Bruxelles, Belgium.

E-mail: ecketr@ccl.org

The authors would like to thank Felix Brodbeck and Marcus Dickson for their helpful comments on an earlier version of this manuscript.

---

© 2009 Psychology Press, an imprint of the Taylor & Francis Group, an Informa business

<http://www.psypress.com/ejwop>

DOI: 10.1080/13594320802678414

interpreting the importance and meaning of self–observer rating discrepancies in 360-degree instruments.

**Keywords:** 360-degree feedback; Self–observer discrepancy; Cultural differences.

Feedback is one of the most important elements for learning and development. In current leadership development initiatives, 360-degree feedback from a variety of rater sources (i.e., boss, peers, direct reports, customers) is a popular method for developmental feedback to managers. This type of feedback is an important mechanism to give managers a sense of how they are perceived by others around them. Feedback information from subordinates, bosses, peers, suppliers, vendors, and other groups highlights individual and organizational strengths and weaknesses (Dalessio, 1998; London & Smither, 1995; Morgeson, Mumford & Campion, 2005; Tornow, 1993; Yammarino, 2003).

Previous research has indicated that a gap or discrepancy exists (i.e., a difference, dissimilarity, disagreement, or incongruity) between self- and observer ratings on the same constructs or dimensions of 360-degree feedback instruments (Brutus, Fleenor & McCauley, 1999; Morgeson et al., 2005; Mount, Judge, Scullen, Sytsma & Hezlett, 1998; Ostroff, Atwater & Feinberg, 2004; Sala, 2003). It is generally argued that such discrepancies have a negative impact for the manager as they indicate inaccurate self-perceptions, low self-awareness, or deficiencies in other soft skills (Kulas & Finkelstein, 2007). However, recent studies (Atwater, Waldman, Ostroff, Robie & Johnson, 2005; Gentry, Hannum, Ekelund & de Jong, 2007) indicate that such interpretation of discrepancies might be too hasty, as discrepancies are also influenced by systematic variables other than managers themselves. In order to interpret discrepancies correctly, it is important to know what underlying factors may be antecedents to such discrepancies.

The current article addresses this issue by examining cultural influences on the self–observer rating discrepancies of leadership skills across a multitude of cultures.

## REASONS FOR RATING DISCREPANCIES

The forms and consequences of rating discrepancies are well explored. Atwater and Yammarino's (1997) model distinguishes the cases of overrating (self > observer), underrating (self < observer), and agreement (self = observer, either in high or in low ratings). Higher disagreement between self- and observer ratings is related to lower performance (Ostroff et al., 2004) and a lower motivation to improve future behaviour, as well as less actual improvement (Atwater & Brett, 2005).

The background and reason for such rating discrepancies, however, has received less attention. The most widely held assumption is that rating discrepancies are an expression of low self-awareness, putting the “blame” for the discrepancy on the manager. However, the validity of interpreting rating discrepancies as low self-awareness is doubted. In their meta-analysis of rating discrepancies, Harris and Schaubroeck (1988) propose three other reasons for rating discrepancy.

First, a manager overrating him/herself could be the result of an inflated self-view or a self-enhancement bias. This explanation resonates with findings that overrating oneself compared to others has the most detrimental effect on performance ratings (Ostroff et al., 2004). Evidence indeed found that some individuals do systematically overrate themselves; however, this was not due to personality differences but to influences such as gender, race, and education (Ostroff et al., 2004). Second, organizational level influences rating discrepancies, because rater groups on different hierarchical levels might have different ideas of what good leadership actually is. Evidence shows that managers on higher levels have bigger rating discrepancies than managers in lower levels (Gentry et al., 2007; Sala, 2003). Third, different rater groups may have different observational opportunities for forming a perception of a target manager. Assuming that leadership behaviour differs according to whom the manager interacts with, the baseline information available to different rater groups is inconsistent. Interestingly, the reasons proposed by Harris and Schaubroeck question the interpretation of rating discrepancies as expressing low self-awareness because they refer to influences only partly within the manager’s control.

Another antecedent that needs attention may be culture of origin. The earlier mentioned research and theory focused on the individual and the organizational level as influence factors. However, as the use of 360-degree instruments increases globally, the level of societal culture becomes more and more important and it would be interesting to know whether cultural values affect such ratings (Brutus, Leslie & McDonald-Mann, 2001). Most recently, Gentry et al. (2007) examined 360-degree ratings of managerial derailment behaviours, and found a culture-level difference: The self-observer rating discrepancy was higher (i.e., bigger or wider) for US-American managers than for Europeans. This suggests that cultural-level determinants, such as values, norms, and beliefs, have an impact on self-observer rating discrepancies.

Taken together, past research shows that multiple reasons for self-observer rating discrepancies exist. In order to make sense of such discrepancies and identify useful avenues for development, managers need information to differentiate discrepancies that are indicative of low self-awareness from those that are due to influences outside their control, such as culture. To understand more about the influence of culture on the

self-observer rating discrepancy, a multilevel approach to 360-degree research is needed that examines culture systematically. Therefore, we examine in the current study how the rating discrepancies between self- and observers on three core leadership skills in a 360-degree feedback instrument called BENCHMARKS<sup>®1</sup> (Lombardo, McCauley, McDonald-Mann & Leslie, 1999) are influenced by cultural values. Although a complete discussion of this issue would integrate all dimensions of cultural value orientations, sources of feedback, and an array of managerial leadership skills, such an effort would be beyond the scope of a single article. Thus, we focus on how power distance, as one exemplary cultural value dimension, might predict rating discrepancies in three leadership skills (decisiveness, leading employees, and composure).

### CULTURAL INFLUENCES ON RATING DISCREPANCIES

In order to examine the influence of cultural values, it is necessary to take a closer look at the cognitive processes in leadership perception. Observers' perception of leaders is based two processes (Lord & Maher, 1991): On the one hand, leadership can be inferred from certain outcomes for which a leader is seen as responsible (inference-based perception); on the other hand, leaders are recognized by comparing their behaviour and characteristics with one's implicit theories about leadership in general (recognition-based perception). In both cases, culture can play a role: Cultural norms on workplace behaviour impact on the amount and type of interaction that various observers have with a leader, resulting in varying opportunities to observe leadership behaviour as well as leadership outcomes. This leads to cultural differences in leadership perception processes (see Ensari & Murphy, 2003, for a detailed study of these differences). Moreover, cultural values also play a role in that they determine observers' implicit theories of leadership (ILTs). As various studies have shown, ILTs vary widely across cultures and covary systematically with cultural values (Brodbeck et al., 2000; Gerstner & Day, 1994). Thus, the same leaders might be perceived as bad, average, or good leaders depending on the culturally determined ILTs of their observers (Schyns, 2006).

What do these cultural influences mean for the discrepancy between self- and observer ratings in 360-degree feedback? Broadly, it can be expected that the magnitude of this discrepancy differs between cultures in which social interaction with the leaders allows observers a very close look at their behaviour, and cultures where leadership is a more "distant" process and

<sup>1</sup>BENCHMARKS<sup>®</sup> is a registered trademark of the Center for Creative Leadership.

observers need to infer leaders' competence from their leadership outcomes, such as group performance. In cultures where observers have a more distant perspective on leaders, the discrepancy between self- and observer ratings will likely be larger.

Various dimensions of cultural values can be involved in this: Stark, Hanson, and Thomas (1998) suggested the dimensions of individualism/collectivism, power distance, and gender egalitarianism as most likely to influence on the scores in 360-degree feedback, due to their impact on the broader social interactions in organizations. However, they did not proceed with empirical tests for these propositions. The current study fills this gap by examining the effects of power distance on rating discrepancy. We chose this culture dimension from the GLOBE study (House, Hanges, Javidan, Dorfman & Gupta, 2004), because it is directly related to the perception of leadership and has been shown to have a great influence on interpersonal interaction at work (e.g., Badjo & Dickson, 2001; Feldman & Bolino, 1999; van der Vegt, van de Vliert & Huang, 2005).

### Power distance

The cultural value dimension of power distance refers to "the extent to which a community accepts and endorses authority, power differences, and status privileges" (Carl, Gupta & Javidan, 2004, p. 513). In high power distance cultures, more emphasis is given to managers' hierarchical status in the organization. Managerial behaviour in these cultures will differ widely depending on the relative status of their interaction partner—boss, peers, or subordinates. The impressions of bosses, peers, and subordinates about the manager's performance will be formed on the basis of very different information, resulting from different observation possibilities, and thus will likely differ from each other more widely than in low power distance cultures, where managerial behaviour will not depend as much on the status of the interaction partner. Specifically, we expect this discrepancy to be expressed in the perception of three leadership skills: decision making, leading employees, and managerial composure.

### Decision making

Decision making is a core leadership task that varies in its style across cultures. As Hofstede (1984) described, one aspect of cultural power distance relates to the preferred style of managerial decision making: High power distance cultures prefer decision making styles that are autonomous or autocratic, and which rarely include others. Low power distance cultures, on the other hand, are characterized by more participative (consultative and democratic) decision-making styles. Managers in low power distance

cultures are thus more likely to include others, such as their bosses, peers and subordinates, into their decision making, thereby giving them better opportunities to observe the manager in decision making situations. Observers are also more likely to have shared opportunities for observation, e.g., in round-table meetings. Based on this observational advantage, we expect that the discrepancy between bosses, peers, and subordinates compared to self-perception decision making in low power distance cultures will be smaller than in high power distance cultures.

**Hypothesis 1:** In high power distance cultures, the self–subordinate, self–boss, and self–peer rating discrepancy on decisiveness is higher than in low power distance cultures.

### Leading employees

We also expect that power distance should relate to the discrepancy between self and observers regarding the manager's skill in leading employees. According to Dansereau, Graen, and Haga (1975), the tasks of a manager can be divided into supervision and leadership. Supervision is more aligned with a cultural orientation towards high power distance (Carl et al., 2004). In environments where this style prevails, there is little rapport between managers and subordinates. The relationships between managers and their subordinates, as well as their bosses, are likely less intense in high power distance cultures than in low power distance cultures (Sullivan, Mitchell & Uhl-Bien, 2003). This should have a negative effect on the congruence of their perception about the manager's skills in leading employees.

Besides the strength of the leadership relationship, feedback behaviour is another factor that could ensure agreement between managers and observers in this area. In high power distance cultures, feedback is given along the vertical lines of hierarchy in organization and is largely unidirectional, top down (Sully de Luque & Sommer, 2000). In these cultures, feedback is given and sought less frequently bottom up, i.e., from subordinates to a manager, or from a manager to higher level manager (Schermerhorn & Bond, 1997). Thus, in low power distance cultures, managers and their bosses (or subordinates) have less bidirectional communication and thus are less likely to be in unison in their perception of a manager's skills of leading employees. Taken together, we hypothesize:

**Hypothesis 2:** In high power distance cultures, the self–subordinate and self–boss rating discrepancy on leading employees is higher than in low power distance cultures.



## Managerial composure

Managerial composure refers to a manager's skill to keep emotions under control, not blame or abuse others, and maintain calmness in times of crisis. As power distance is a cultural dimension that emphasizes behaviours to ensure and enhance power and social status, we also expect that it should influence the discrepancy between self and others in the perception of managerial composure. For a leadership style mostly characterized by supervisory (rather than leadership) behaviours, managers need to show high composure and ensure they maintain their social status, especially towards their subordinates. In their relationship with subordinates, for example, managers can be expected to show less affect, responsiveness, and equivalence in their social interactions (Sullivan et al., 2003). Moreover, their behavioural leeway is smaller, defined mostly by norms and expectations towards their social roles rather than them as individual persons. In low power distance cultures, however, managers focus less on supervisory tasks and engage more in direct leadership behaviours. In those leadership relationships, close leadership relationships are built and trust is established, which in turn enable managers and subordinates to be more personal, emotional, and approachable than in high power distance cultures. Thus, we expect that managers and subordinates in low power distance cultures are more aligned in their perception of managerial composure than those in high power distance cultures.

**Hypothesis 3:** In high power distance cultures, the self-subordinate rating discrepancy on composure is higher than in low power distance cultures.

## METHOD

### Participants and procedure

Data were obtained via the use of the BENCHMARKS<sup>®</sup> assessment tool in leadership development feedback exercises. Managers who took part in a leadership development process between October 2000 and May 2006 provided self-ratings on this instrument and selected observers (one boss, a minimum of three subordinates, and minimum of three peers) to provide ratings on the same tool. Because culture (i.e., country) was an integral part of the analysis, we only selected cases where self and observers were native to and currently working in the same country. We also set a minimum of 15 self-raters from each country. The final sample totalled 4019 managers (self-raters) from 31 countries. The average age of these managers was 41.6 years, 71.8% were male, 83.3% had at least a Bachelor's level of education, 69.5% were middle or upper-middle level managers. Table 1 displays the number of managers from each country.



TABLE 1  
Number of managers within countries studied

<i>Country</i>	<i>n</i>
Argentina	30
Australia	170
Brazil	60
Canada	815
Denmark	29
Finland	15
France	110
Germany	133
Hong Kong	21
India	100
Indonesia	36
Ireland	66
Italy	34
Japan	16
Malaysia	21
Mexico	128
Netherlands	190
New Zealand	77
Philippines	67
Poland	31
Portugal	25
Russia	18
Singapore	199
South Korea	31
Spain	177
Sweden	31
Switzerland	39
Thailand	23
Turkey	28
United Kingdom	484
United States*	815

\*In order to keep sample size differences within proportions, the sample from the United States was a random sample of 815 managers selected from a much larger country data pool.

Measures

BENCHMARKS<sup>®</sup> is a multisource feedback instrument, accumulating ratings from the self, direct report, peer, and boss perspective (Lombardo & McCauley, 1994). It is a well-validated and reviewed 360-degree feedback instrument (Carty, 2003; Center for Creative Leadership, 2004; Douglas, 2003; McCauley, Lombardo & Usher, 1989; Spangler, 2003; Zedeck, 1995) and has also been used in prior research (e.g., Atwater, Ostroff, Yammarino

& Fleenor, 1998; Brutus, Fleenor & London, 1998; Brutus et al., 1999; Conway, 2000; Fleenor, McCauley & Brutus, 1996; Gentry et al., 2007). Raters used a 5-point Likert-type scale determining the extent to which a manager displays a leadership skill with 1 = "Not at all" to 5 = "To a very great extent". All items were identical for managers and their observers, only the item referent changed. We selected the three specific leadership dimensions of decisiveness, leading employees, and managerial composure, as the analyses of Braddy, Fleenor, and Campbell (2008) have shown that these scales showed good discriminative power on item and scale level, as well as sufficient scale reliabilities, to be considered separate leadership skills.

*Decisiveness.* This scale contains four items and refers to a preference for action rather than thinking, taking calculated risks, and being a quick decision maker when necessary ( $\alpha = .78$ ).

*Leading employees.* This scale contains 11 items about delegating and broadening the skills of direct reports, being fair and patient, and setting clear performance goals ( $\alpha = .90$ ).

*Managerial composure.* This scale contains four items that relate to tendencies of blaming and abusing others, not being arrogant, cynical or moody, and coping with situations beyond one's personal span of control ( $\alpha = .74$ ).

*Cultural values.* We used the GLOBE values ("should be") scores to assess culture-level power distance (see House et al., 2004, for complete tables of these values). Similar value measures are available also from the Hofstede studies, but these measures have been critiqued for their low reliability (Spector, Cooper & Sparks, 2001) and questions of their validity, as they have shown shortcomings in replicating consistent patterns in later studies (Merritt, 2000). This scale measured power distance values with five items ( $\alpha = .74$  across the GLOBE cultures).

## Aggregation issues

An average of 4.14 direct reports and 3.92 peers rated each manager. To justify aggregation of ratings across observers, we computed ICC(1) and ICC(2) (Bliese, 2000), as well as  $r^*_{WG(J)}$  (Lindell, Brandt & Whitney, 1999). These values are given in Table 2 and indicate suitable requirements to aggregate multisource feedback ratings (Greguras & Robie, 1998; James, 1988) and are similar to values reported in other multisource studies (Atwater et al., 1998; Fleenor et al., 1996; Ostroff et al., 2004), thus we

TABLE 2  
Aggregation statistics, means, standard deviations, and intercorrelations of study variables

	<i>ICC(1)</i>	<i>ICC(2)</i>	<i>r*<sub>wG(j)</sub></i>	<i>Mean</i>	<i>SD</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>	<i>11</i>	<i>12</i>
1. Power distance				2.70	0.30												
2. Decisiveness—self				3.83	0.57	-.39											
3. Leading employees—self				3.83	0.45	-.20	.56										
4. Composure—self				3.87	0.51	-.21	.41	.44									
5. Decisiveness—boss				3.93	0.60	-.23	.26	.16	.06								
6. Leading employees—boss				3.80	0.52	-.22	.15	.25	.10	.57							
7. Composure—boss				3.97	0.60	-.26	.01	.06	.18	.46	.51						
8. Decisiveness—peers	.25	.56	.78	3.86	0.43	-.26	.32	.19	.09	.41	.30	.15					
9. Leading employees—peers	.25	.57	.77	3.69	0.39	-.13	.18	.28	.15	.25	.41	.24	.64				
10. Composure—peers	.26	.57	.75	3.83	0.45	-.27	.01	.05	.25	.13	.23	.41	.44	.60			
11. Decisiveness—subordinates	.27	.60	.76	3.93	0.45	-.13	.29	.21	.08	.31	.25	.10	.45	.34	.16		
12. Leading employees—subordinates	.26	.59	.71	3.79	0.44	.11	.17	.30	.12	.19	.32	.14	.29	.43	.22	.68	
13. Composure—subordinates	.27	.61	.72	3.91	0.47	-.11	.03	.11	.24	.10	.22	.29	.18	.31	.42	.53	.64

*N* = 4019 for all variables except power distance, where statistics are shown at the country level, thus *N* = 31. Culture-level correlations with  $|r| > .35$  are significant with  $p < .05$ ; all individual-level correlations are significant with  $p < .01$ . All *F*s for *ICC*(1) > 2.25.

proceeded the analysis with aggregated scores across direct reports and peers.

### Analytic procedure

The effects predicted by the hypotheses are all cross-level effects, in that the predictor variables (culture) are all measured at a higher level (country), while the outcome variables (discrepancy between raters) are at the individual manager level. Therefore it was necessary to employ a multilevel modelling framework (e.g., Bliese, 2002). The predicted variables were modelled as a discrepancy between self and observer ratings rather than simple difference scores. The problems with using difference scores for this type of variable are many: Because difference scores combine two distinct pieces of information into one, they cannot be unambiguously interpreted, lead to decreased reliability, put what is supposed to be measured with a multivariate model into a univariate framework, and consequently, can provide misleading results (Edwards, 1994, 1995), particularly in cases where measures are positively related (Johns, 1981). Therefore we used Edwards' (1995) framework for studying discrepancy as an outcome variable, which involves studying the simultaneous effects on ratings from different sources and determining whether these effects differ significantly. This procedure is based on multiple regression with concurrent dependent variables (or MANOVA), and examines the extent to which the effect of predictors on outcomes is different—in our case, examining the differential effect of power distance on ratings of different observers. This method is described in detail in Edwards (1995). However, as we needed to do this within a multilevel framework, we used Mplus (Muthén & Muthén, 2006) to run multilevel structural equation models (with manifest variables) in analogy to polynomial regression, which allowed the testing of differences via the imposing of constraints on parameters. All forms of ratings (self, boss, peers, subordinates) were modelled simultaneously as outcome variables on the culture variable of power distance. We set constraints to test the equivalence of the effects of power distance on each rating. Due to the use of the maximum likelihood ratio (MLR) estimator, we used Satorra's (2000) test for adjusted chi-squared values.

## RESULTS

Table 2 shows aggregation statistics, means, standard deviations and intercorrelations of all study variables. Intercorrelations are given at the individual (manager) level, except for those involving the culture variables, which are given at the country level.

Hypotheses 1–3 were tested using the analytic procedure described earlier. A summary of the results from these tests is shown in Table 3,

TABLE 3  
Summary of results for Hypotheses 1 to 3

Power distance	Effect size for ...			
	Self	Boss	Peers	Subordinates
H1: Decisiveness	−.23	−.16	−.16	−.07*
H2: Leading employees	−.16	−.12	—	.00*
H3: Composure	−.08	—	—	—

Figures in table are unstandardized regression weights (paths) in the unconstrained model. \*indicates a difference from the effects on “Self” with  $p < .05$ .

showing partial support for our hypotheses. However, to understand the effects fully, it was necessary to plot them: These plots are shown in Figures 1, 2, and 3. The values depicted for the high and low end of power distance resemble scores of  $\pm 1$  SD from the grand mean. For the sake of completeness, all four rater sources are shown in each figure.

Hypothesis 1 was only partially supported, as the only significant rating discrepancy in decisiveness emerged between self and subordinates, and in the hypothesized direction. Figure 1 shows that self-ratings of decisiveness decrease considerably in cultures of high power distance, and thus these

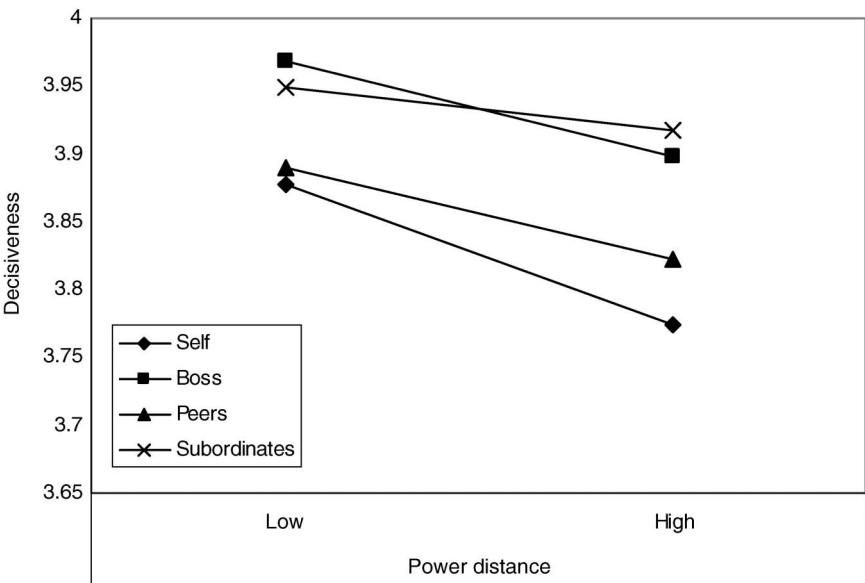


Figure 1. Differential effects of power distance on ratings of decisiveness.

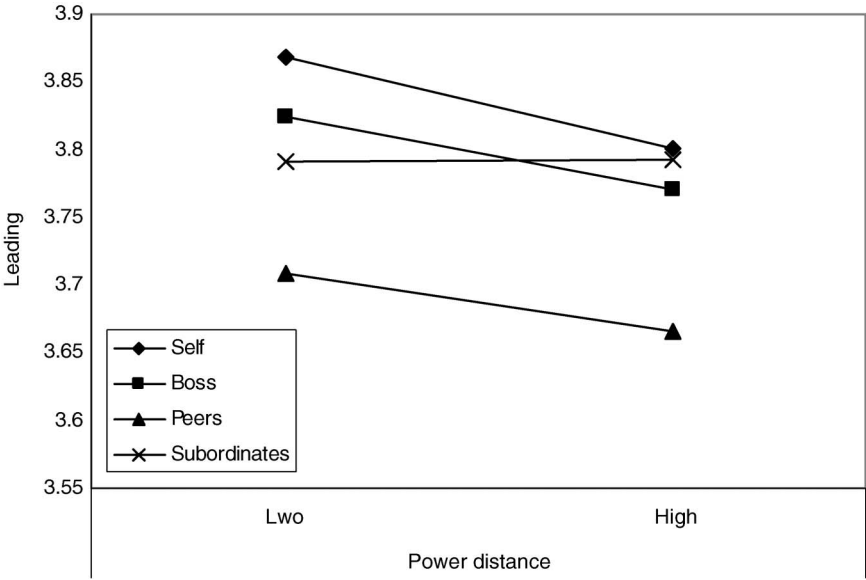


Figure 2. Differential effects of power distance on ratings of leading employees.

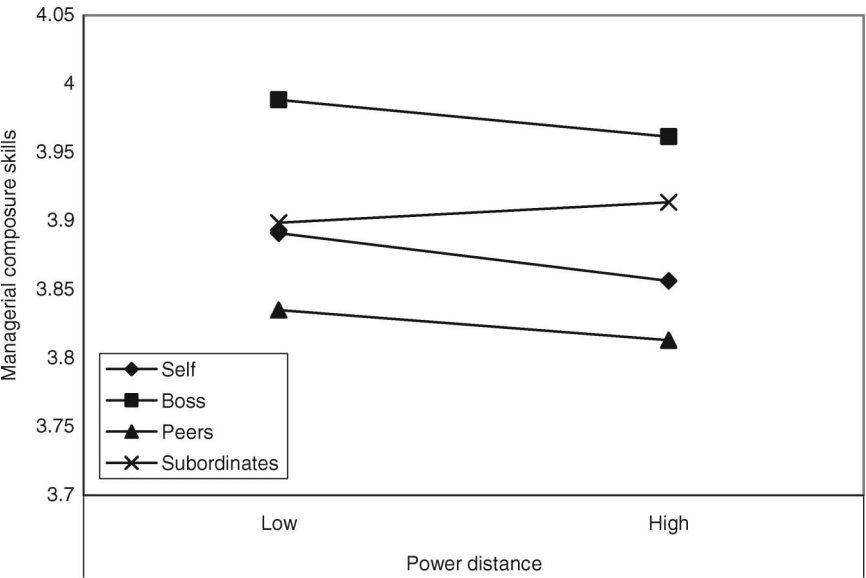


Figure 3. Differential effects of power distance on ratings of composure.

cultures show a higher discrepancy between self and observer ratings, most notably, subordinates. Hypothesis 2 regarding leading employees was not supported, as the only significant discrepancy, which again emerged between self- and subordinate ratings, widened with lower power distance scores, which was contrary to what we hypothesized (mostly because of a decrease in self-ratings, see Figure 2). Hypothesis 3 concerning managerial composure was supported as the discrepancy between self- and subordinate ratings widened with higher power distance (both due to increase in subordinate ratings and decrease in self ratings, see Figure 3).

## DISCUSSION

In this study we examined the effect of power distance on perception discrepancy between managers, their bosses, peers, and subordinates, in three core leadership skills: Decisiveness, leading employees, and composure. Using multilevel modelling, we assessed the discrepancy of each kind of observer rating compared to self-ratings.

Our results did not support all our hypotheses, but in sum they strongly suggest that cultural values have a systematic effect on rating discrepancies between self and observers. As expected, we found that in high power distance cultures, the discrepancy between self and subordinate ratings about a manager's decisiveness and composure was higher than in low power distance cultures. We also found that the discrepancy between self- and subordinate ratings about a manager's skills in leading employees was lower in high-power distance culture, contrary to our hypothesis.

Overall, these results show that power distance indeed influences rating discrepancies between self and observers, especially between subordinate ratings, of leadership skills. How a manager's skills are perceived depends on the opportunities different observers have to observe the leader. These opportunities, in turn, depend on the quality and quantity of interaction with the manager, which is determined by cultural norms and values.

Although we examined the impact of cultural values only on selected discrepancies and three specific leadership skills, it can be assumed that cultural values also influence the perception of managers and their observers in other areas of leadership. Traditionally, a discrepancy between self- and observer ratings is regarded as a negative indication of the leadership relationship, or even as a lack of self-awareness with respect to the manager (Kulas & Finkelstein, 2007). Our results question the applicability and generalizability of such conceptualizations once a cross-cultural perspective is considered. Clearly, more research is necessary to examine the impact of rating agreement or discrepancy across cultures, ideally using theoretical frameworks such as dimensions of cultural values or practices.



## Limitations

Our analyses were carried out with appropriate consideration to the multilevel nature of the effects in question. Limitations towards the external validity of our results, however, arise from the circumstance that most of the participants in this archival data study could be considered as having high leadership potential by their company, as they participated in 360-degree feedback for the purpose of their own development, rather than for assessment purposes. Such a sample of managers might not be representative of the population of managers at large.

Another limitation might stem from language: All participants received the 360-degree feedback tool in English, even if this was not their mother tongue. Thus, there is a certain risk that language might have influenced results, although it would be hard to understand how this might have led to the systematic differences that we found. Moreover, as reliabilities of the scales were high across all cultures, the impact of such a bias on our results should be rather small.

Finally, it should be noted that the country-level sample sizes differed and sometimes were rather small. This could have led to increased variability in the responses of managers within these countries, in turn leading to systematic underestimation of culture-level influences and overly conservative hypothesis tests.

## Implications for research

Research should now consider culture, along with individual and organizational level variables in tandem, as well as the relationship between manager and observer, as influences for the pattern of self and observer ratings in 360-degree feedback. Examining the interactive effects of individual, organizational, and cultural variables as well as the relationship between managers and observers may give great insight into why rating discrepancies exist.

Specifically, we see three main directions for future research. First, it would be important to look not only at the magnitude of perception discrepancies, but also at their direction, examining whether there might be culturally determined patterns in the way managers systematically over- or underrate themselves compared to their observers.

Second, our study implies that the relationship of rating discrepancies with managerial outcomes, such as leadership performance, employee satisfaction, or promotion, will likely differ across cultures. If discrepancies are not based on shortcomings from the manager (e.g., a lack of self-awareness), then why should they have negative consequences? Although research in Western cultures shows that higher discrepancies are related to

negative career outcomes, such a relationship cannot simply be assumed to hold true in other cultures. First evidence for this notion can be found in Atwater et al. (2005), showing less importance of rater agreement in European countries than in the US, as well as significant variation between European countries. Similar studies including systematic differences in values, development, or other country characteristics need to be conducted to help us judge the predictive value of rating agreement overall.

Third, cases of expatriate leadership or multicultural leadership, where managers and observers have different cultural backgrounds, should also be examined. In these cases, the cultural influence on rating discrepancies does not only stem from value differences but also from differences in culturally endorsed implicit leadership theories, which determine the course of leadership perception (Den Hartog et al., 1999). In order to really understand the role of culture in 360-degree feedback ratings globally, these results should then be integrated into a comprehensive model about the influence of leadership perception by multiple stakeholders in multiple cultures.

### Implications for practice

When managers receive their 360-degree feedback and look at the discrepancies between self- and observer ratings, they are often confused. Managers may not only be at a loss about which information to trust (their own thoughts, or the perception of their observers), but also may not understand what the reasons for such discrepant information might be. Our study demonstrates that one reason for such discrepancy lies within cultural values rather than managers themselves. This insight is of high practical relevance as it can help prevent misinterpretation of rating discrepancies and is particularly relevant as many organizations are using 360-degree feedback around the world (Atwater et al., 2005; Brutus et al., 2001). It raises question to the traditional interpretation of rating discrepancies of low self-awareness and the implication of such discrepancies for future leadership development.

The interpretation of rating discrepancies is based on results about the negative impact of overrating and highly discrepant ratings on managerial performance and derailment for managers in the United States (e.g., Atwater et al., 1998; van Velsor & Leslie, 1991). Managers who do not achieve congruence between their self-perception and their impact on others are likely to face negative outcomes for their career (Wohlers & London, 1989; Yammarino & Atwater, 1993). If discrepancies are in fact driven by cultural values, such interpretations are not appropriate and managers will need to be counselled differently about what to do with their feedback. Indeed, we assume that managerial attempts to reduce discrepancies that are based on cultural issues will likely not be fruitful—nor necessarily appreciated. In order to create the conditions for lower rating discrepancy,

managers could find themselves in a situation where they would have to behave contrary to the culturally endorsed expectations about good leadership. Such a violation of expectations would likely result in a negative reaction of others around them and not be conducive to their overall leadership effectiveness.

Our results also question whether high rating discrepancies should be continued to be regarded as something negative that needs to be improved or a deficiency that needs development. Attempts to reduce such discrepancies without questioning their origin is oversimplified and will not always lead to satisfying results. Shipper, Hoffman, and Rotondo (2007) have already found that US managers who engage in 360-degree feedback for their own development and take it repeatedly have a higher decrease in their perception discrepancy than managers from other countries. If, as we have shown, these discrepancies are due to contextual influences rather than individual leadership capabilities, such findings about cultural variations in the longitudinal development of discrepancies are not surprising.

Lastly, our findings have implications for coaches of cross-cultural managers: They can use our findings and purposefully include culture-level information and advice to cross-cultural managers when going through their 360-degree feedback with them.

To sum up, our study shows that cultural values systematically influence how managers and others around them perceive various leadership skills. This has wide-reaching implications of the interpretation of rating discrepancies in 360-degree feedback. In order to fully understand these discrepancies and their practical meaning, further systematic research into the unique and interactive effects of culture on leadership perception from the perspectives of self and observers is needed.

## REFERENCES

- Atwater, L. E., & Brett, J. F. (2005). Antecedents and consequences of reactions to developmental 360° feedback. *Journal of Vocational Behavior*, 66, 532–548.
- Atwater, L. E., Ostroff, C., Yammarino, F. J., & Fleenor, J. W. (1998). Self-other agreement: Does it really matter? *Personnel Psychology*, 51(3), 577–598.
- Atwater, L., Waldman, D., Ostroff, C., Robie, C., & Johnson, K. M. (2005). Self-other agreement: Comparing its relationship with performance in the US and Europe. *International Journal of Selection and Assessment*, 13, 25–40.
- Atwater, L. E., & Yammarino, F. (1997). Self-other rating agreement: A review and model. *Research in Personnel and Human Resources Management*, 15, 141–164.
- Badjo, L. M., & Dickson, M. W. (2001). Perception of organizational culture and women's advancement in organizations: A cross-cultural examination. *Sex Roles*, 45(5/6), 399–414.
- Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analyses. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 349–381). San Francisco: Jossey-Bass.

- Bliese, P. D. (2002). Multilevel random coefficient modeling in organizational research: Examples using SAS and S-PLUS. In F. Drasgow & N. Schmitt (Eds.), *Measuring and analyzing behavior in organizations: Advances in measurement and data analysis* (pp. 401–445). San Francisco: Jossey-Bass.
- Braddy, P. W., Fleenor, J. W., & Campbell, M. (2008). *Methods for reducing the length of 360-degree instruments*. Paper presented at the 23rd annual conference of the Society for Industrial and Organizational Psychology, San Francisco.
- Brodbeck, F. C. (2000). Cultural variation of leadership prototypes across 22 European countries. *Journal of Occupational and Organizational Psychology*, 73(1), 1–29.
- Brutus, S., Fleenor, J. W., & London, M. (1998). Does 360-degree feedback work in different industries? A between-industry comparison of the reliability and validity of multi-source performance ratings. *Journal of Management Development*, 17, 177–190.
- Brutus, S., Fleenor, J. W., & McCauley, C. D. (1999). Demographic and personality predictors of congruence in multi-source ratings. *Journal of Management Development*, 18, 417–435.
- Brutus, S., Leslie, J. B., & McDonald-Mann, D. (2001). Cross-cultural issues in multisource feedback. In D. Bracken, C. Timmreck, & A. Church (Eds.), *Handbook of multisource feedback* (pp. 433–446). San Francisco: Jossey-Bass.
- Carl, D., Gupta, V., & Javidan, M. (2004). Power distance. In R. J. House, P. J. Hanges, M. Javidan, P. W. Dorfman, & V. Gupta (Eds.), *Culture, leadership, and organizations: The GLOBE study of 62 societies* (pp. 513–563). Thousand Oaks, CA: Sage.
- Carty, H. M. (2003). Review of BENCHMARKS® [revised]. In B. S. Plake, J. Impara, & R. A. Spies (Eds.), *The fifteenth mental measurements yearbook* (pp. 123–124). Lincoln, NE: Buros Institute of Mental Measurements.
- Center for Creative Leadership. (2004). *BENCHMARKS® facilitator's manual*. Greensboro, NC: Center for Creative Leadership.
- Conway, J. M. (2000). Managerial performance development constructs and personality correlates. *Human Performance*, 13, 23–46.
- Dalessio, A. T. (1998). Using multi-source feedback for employee development and personnel decisions. In J. W. Smither (Ed.), *Performance appraisal: State-of-the-art in practice* (pp. 278–330). San Francisco: Jossey-Bass.
- Dansereau, F., Graen, G., & Haga, W. J. (1975). A vertical dyad linkage approach to leadership within formal organizations: A longitudinal investigation of the role making process. *Organizational Behavior and Human Performance*, 13, 46–78.
- Den Hartog, D. N., House, R. J., Hanges, P. J., Ruiz-Quintanilla, S. A., Dorfman, P. W., & GLOBE co-authors. (1999). Culture specific and cross-culturally generalizable implicit leadership theories: Are attributes of charismatic/transformational leadership universally endorsed? *Leadership Quarterly*, 10(2), 219–256.
- Douglas, C. A. (2003). *Key events and lessons for managers in a diverse workforce: A report on research and findings*. Greensboro, NC: Center for Creative Leadership.
- Edwards, J. R. (1994). The study of congruence in organizational behavior research: Critique and proposed alternative. *Organizational Behavior and Human Decision Processes*, 58, 51–100.
- Edwards, J. R. (1995). Alternatives to difference scores as dependent variables in the study of congruence in organizational research. *Organizational Behavior and Human Decision Processes*, 64, 307–324.
- Ensari, N., & Murphy, S. E. (2003). Cross-cultural variations in leadership perceptions and attribution of charisma to the leader. *Organizational Behavior and Human Decision Processes*, 92(1–2), 52–66.
- Feldman, D. C., & Bolino, M. C. (1999). The impact of on-site mentoring on expatriate socialisation: A structural equation modelling approach. *International Journal of Human Resource Management*, 10(1), 54–71.

- Fleenor, J. W., McCauley, C. D., & Brutus, S. (1996). Self-other rating agreement and leader effectiveness. *Leadership Quarterly*, 7, 487–506.
- Gentry, W. A., Hannum, K. M., Ekelund, B. Z., & de Jong, A. (2007). A study of the discrepancy between self- and observer-ratings on managerial derailment characteristics of European managers. *European Journal of Work and Organizational Psychology*, 16, 295–325.
- Gerstner, C. R., & Day, D. V. (1994). Cross-cultural comparison of leadership prototypes. *Leadership Quarterly*, 5(2), 121–134.
- Greguras, G. J., & Robie, C. (1998). A new look at within-source interrater reliability of 360-degree feedback ratings. *Journal of Applied Psychology*, 83, 960–968.
- Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, 41, 43–61.
- Hofstede, G. (1984). *Culture's consequences: International differences in work-related values*, Vol. 2. Newbury Park, CA: Sage.
- House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta V. (Eds.). (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Thousand Oaks, CA: Sage.
- James, L. R. (1988). Organizational climate: Another look at a potentially important construct. In S. G. Cole & R. G. Demaree (Eds.), *Applications of interactionist psychology: Essays in honor of Saul B. Sells* (pp. 253–282). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Johns, G. (1981). Difference score measures of organizational behavior variables: A critique. *Organizational Behavior and Human Performance*, 27, 443–463.
- Kulas, J. T., & Finkelstein, L. M. (2007). Content and reliability of discrepancy-defined self-awareness in multisource feedback. *Organizational Research Methods*, 10(3), 502–522.
- Lindell, M. K., Brandt, C. J., & Whitney, D. J. (1999). A revised index of interrater agreement for multi-item ratings of a single target. *Applied Psychological Measurement*, 23, 127–135.
- Lombardo, M. M., & McCauley, C. D. (1994). *BENCHMARKS®: A manual and trainer's guide*. Greensboro, NC: Center for Creative Leadership.
- Lombardo, M. M., McCauley, C. D., McDonald-Mann, D., & Leslie, J. B. (1999). *BENCHMARKS® developmental reference points*. Greensboro, NC: Center for Creative Leadership.
- London, M., & Smither, J. W. (1995). Can multi-source feedback change self-evaluations, skill development, and performance? Theory-based applications and directions for research. *Personnel Psychology*, 48, 375–390.
- Lord, R. G., & Maher, K. J. (1991). *Leadership and information processing: Linking perceptions and performance*. London: Routledge.
- McCauley, C., Lombardo, M., & Usher, C. (1989). Diagnosing management development needs: An instrument based on how managers develop. *Journal of Management*, 15, 389–403.
- Merritt, A. C. (2000). Culture in the cockpit: Do Hofstede's dimensions replicate? *Journal of Cross-Cultural Psychology*, 31, 283–301.
- Morgeson, F. P., Mumford, T. V., & Campion, C. A. (2005). Coming full circle: Using research and practice to address 27 questions about 360-degree feedback programs. *Consulting Psychology Journal: Practice and Research*, 57, 196–209.
- Mount, M. K., Judge, T. A., Scullen, T. E., Sytsma, M. R., & Hezlett, S. A. (1998). Trait, rater, and level effects in 360-degree performance ratings. *Personnel Psychology*, 51, 557–576.
- Muthén, L., & Muthén, B. (2006). *Mplus (Version 4.0)* [Computer software]. Los Angeles: Author.
- Ostroff, C., Atwater, L. E., & Feinberg, B. J. (2004). Understanding self-other agreement: A look at rater and ratee characteristics, context, and outcomes. *Personnel Psychology*, 57, 333–375.
- Sala, F. (2003). Executive blind spots: Discrepancies between self- and other-ratings. *Consulting Psychology Journal*, 55, 222–229.

- Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H. Heimans, D. S. G. Pollock, & A. Satorra (Eds.), *Innovations in multivariate statistical analysis* (pp. 233–247). London: Kluwer Academic Publishers.
- Schermerhorn, J. R., & Bond, M. H. (1997). Cross-cultural leadership dynamics in collectivism and high power distance settings. *Leadership and Organization Development Journal*, 18(4), 187–193.
- Schyns, B. (2006). The role of implicit leadership theories in the performance appraisals and promotion recommendations of leaders. *Equal Opportunities International*, 25(3), 188–199.
- Shipper, F., Hoffman, R. C. I., & Rotondo, D. M. (2007). Does the 360 feedback process create actionable knowledge equally across cultures? *Academy of Management Learning and Education*, 6(1), 33–50.
- Spangler, M. (2003). Review of BENCHMARKS® [revised]. In B. S. Plake, J. Impara, & R. A. Spies (Eds.), *The fifteenth mental measurements yearbook* (pp. 124–126). Lincoln, NE: Buros Institute of Mental Measurements.
- Spector, P. E., Cooper, C. L., & Sparks, K. (2001). An international study of the psychometric properties of the Hofstede Values Survey Module 1994: A comparison of individual and country/province level results. *Applied Psychology: An International Review*, 50(2), 269–281.
- Stark, E., Hansen, J., & Thomas, L. T. (1998). *A field study of accuracy, agreement, and outcomes of multi-source feedback: Implications across international boundaries*. Paper presented at the international conference of the Academy of Business and Administrative Sciences.
- Sullivan, D. M., Mitchell, M., & Uhl-Bien, M. (2003). The new conduct of business: How LMX can help capitalize on cultural diversity. In G. Graen (Ed.), *Dealing with diversity* (pp. 183–218). Charlotte, NC: Information Age.
- Sully de Luque, M. F., & Sommer, S. M. (2000). The impact of culture on feedback-seeking behavior: An integrated model and propositions. *Academy of Management Review*, 25, 829–849.
- Tornow, W. W. (1993). Perceptions or reality: Is multi-perspective measurement a means or an end? *Human Resource Management*, 32, 246–264.
- Van der Vegt, G. S., van de Vliert, E., & Huang, X. (2005). Location-level links between diversity and innovative climate depend on national power distance. *Academy of Management Journal*, 48(6), 1171.
- Van Velsor, E., & Leslie, J. B. (1991). *Feedback to managers: Vol. 1. A guide to evaluating multi-rater feedback instruments*. Greensboro, NC: Center for Creative Leadership.
- Wohlers, A., & London, M. (1989). Ratings of managerial characteristics: Evaluation difficulty, co-worker agreement and self-awareness. *Personnel Psychology*, 42, 235–261.
- Yammarino, F. (2003). Modern data analytic techniques for multisource feedback. *Organizational Research Methods*, 6, 6–14.
- Yammarino, F., & Atwater, L. (1993). Understanding self-perception accuracy: Implications for human resource management. *Human Resource Management*, 32(2/3), 231–247.
- Zedeck, S. (1995). Review of BENCHMARKS®. In J. Conoley & J. Impara (Eds.), *The twelfth mental measurements yearbook* (Vol. 1, pp. 128–129). Lincoln, NE: Buros Institute of Mental Measurements.